

Bayesian Analysis for miRNA and mRNA Interactions Using Expression Data

Mingjun Zhong, Rong Liu, Bo Liu
 {zhong,rliu,lbo}@dlut.edu.cn

Department of Biomedical Engineering
 Dalian University of Technology, 116023, PR. China

Abstract

MicroRNAs (miRNAs) are small RNA molecules composed of 19-22 nt playing important regulatory roles in post-transcriptional gene regulation by inhibiting the translation of the mRNA into proteins or otherwise cleaving the target mRNA. Inferring miRNA targets provides useful information for understanding the roles of miRNA involving in biological processes which may result in diagnosing complex diseases. Statistical methodologies of point estimates such as the LASSO algorithm have been proposed to identify the interactions of miRNA and mRNA based on sequence and expression data. In this paper, we propose Bayesian LASSO and non-negative Bayesian LASSO to analyze the interactions between miRNA and mRNA using the expression data. The proposed Bayesian methods explore the posterior distributions for those parameters required in the model depicting the miRNA-mRNA interactions. For comparison purposes, we applied the Least Square Regression (LSR), Ridge Regression (RR), LASSO, non-negative LASSO (nLASSO), and the Bayesian approaches to four public data sets which have the known interaction pairs of miRNA and mRNA. Comparing to the point estimate algorithms, the Bayesian methods are able to infer more known interactions and are more meaningful to provide credible intervals to take into account the uncertainty of the interactions of miRNA and mRNA. The Bayesian approaches are useful for graphing the inferred effects of the miRNAs on the targets by plotting the posterior distributions of those parameters, and while the point estimate algorithm only provides a single estimate for those parameters.

1 Introduction

Mature miRNAs, which are a large family of regulatory RNAs expressed in animals and plants, are processed from primary transcripts in two steps by enzymes of the RNase III endonucleases Drosha in the nucleus and Dicer in the cytoplasm (Lee et al., 2003; Hutvagner, 2005). In cytoplasm, miRNA is then separated into two single strands by helicase enzymes and then one strand is combined with an Argonaute protein to form the RNA-induced silencing complex (RISC), which then mediates to regulate gene expression by base pairing imperfectly to the 3'-untranslated region of target mRNAs to inhibit protein synthesis (Carthew and Sontheimer, 2009).

Recent researches have revealed that there exists more than 1400 human miRNA sequences (Kozomara and Griffiths-Jones, 2011) and it has been reported that miRNA plays important roles in various regulation processes including proliferation, apoptosis, differentiation and development, cellular identity and pathogen-host interactions (Parker and Sheth, 2007; Pillai et al., 2007; Carthew and Sontheimer, 2009). Given the measured sequences of miRNA and mRNA, an important task is to identify the regulation interactions between them. Experimental and sequence-based methods have been proposed to predict targets of miRNA regulations. Examples of those algorithms include miRanda (Betel et al., 2008), TargetScan (Lewis et al., 2005), miRBase (Griffiths-Jones et al., 2008) and PicTar (Krek et al., 2005). These algorithms are mainly based on the experimentally determined rules of miRNA regulation. The most general feature of miRNA regulation is that the 3'-untranslated region of target mRNAs is complementary to the seed region (nucleotides 2-7) of the miRNA (Lewis et al., 2003). It has been shown that computational approaches for combining miRNA and mRNA high-throughout expression data with sequence-base putative predictions can greatly reduce the number of targets of miRNA regulation (Muniatégui et al., 2012b).

Various computational algorithms have been proposed to identify the mRNA and miRNA interactions using high-throughout expression data sets. These algorithms include correlation, mutual information, multiple linear regression, partial least squares, LASSO (Tibshirani, 1996), non-negative LASSO (TaLASSO), variational approximations for linear regression models (GenMiR++), HCTarget, and Bayesian graphical approaches, etc. Muniatégui et al. (2012b) provides a review for these algorithms. Among those proposed methods, GenMiR++ (Huang et al., 2007), HCTarget (Su et al., 2011) and the Bayesian graphical approaches (Stingo et al., 2010) are probabilistic approaches. HCTarget and Bayesian graphical method employ the MCMC schemes for sampling the posterior densities for the required parameters, and while the GenMiR++ employs the variational approach to approximate the densities of the parameters. The Bayesian method of Stingo et al. (2010) is a rather different approach for determining the interactions of mRNA and miRNA, comparing to the linear regression models. Both the LASSO algorithm and the non-negative LASSO (nLASSO) algorithm have been applied to the linear regression models for modelling the miRNA-mRNA interactions (Lu et al., 2011; Muniatégui et al., 2012a). In this paper, we propose to employ the Bayesian LASSO and non-negative Bayesian LASSO for evaluating the miRNA regulations. The Bayesian LASSO algorithm is similar to HCTarget and GenMiR++ where no non-negative constraints are put on the linear regression model parameters. We proposed the non-negative Bayesian LASSO algorithm based on the TaLASSO (Muniatégui et al., 2012a) where the non-negative constraints on the parameters have been shown to be more meaningful in searching for the down regulation effects of miRNA on the mRNA expression. The idea of employing the non-negative constraints is that miRNAs are assumed to down regulate their corresponding mRNA targets. Given the linear regression model, we could employ various

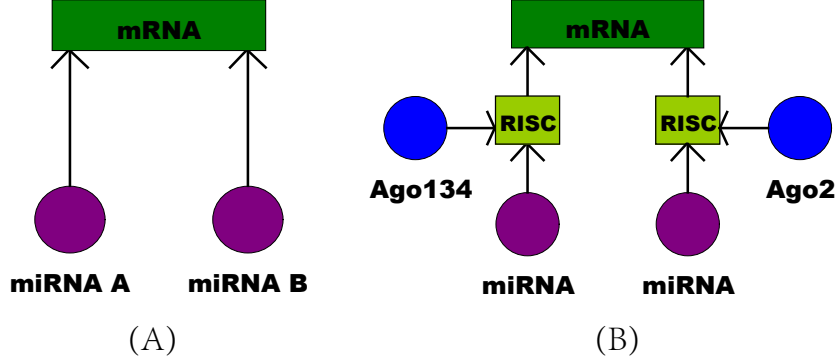


Figure 1: The models for mRNA and miRNA interactions: (A) miRNA directly targets mRNA and (B) the complex RISC composed of miRNA and Argonaute proteins targets mRNA.

algorithms for resolving the model. For this study we have applied the Least Square Regression (LSR), Ridge Regression (RR), Least Absolute Shrinkage and Selection Operator (LASSO), non-negative LASSO (nLASSO), Bayesian LASSO (BLASSO) and non-negative Bayesian LASSO (nBLASSO) to the experimentally validated interactions of four expression data sets. For the point estimate algorithms, we have observed that LSR, RR and LASSO produced the same results and while the nLASSO was able to produce sparse representations for the mRNA-miRNA interactions. The Bayesian approaches of BLASSO and nBLASSO output similar results. Since both the BLASSO and nBLASSO are not able to select variables automatically, we propose to calculate the credible intervals to take into account the prediction uncertainties for the variable selections. The statistical significance is calculated for the credible intervals and then the variables are selected based on this statistical significance given a significance probability α . Thus the numbers of selected variables are dependent on the given significance probability. We concluded that the Bayesian approaches produce more meaningful results, which can measure the uncertainty of the predictions, than the point estimate based algorithms which only produce a single point for the parameters.

2 The Model

Given a set of expression data of miRNA and mRNA, we model the interactions between mRNA and miRNA by the following mathematical model

$$\mathbf{Y} = \mathbf{f}(\mathbf{X}) + \mathbf{E}$$

where $\mathbf{Y} \in \mathcal{R}^{N \times P}$ is the collections of the mRNA expression data where N is the number of tissue samples and P is the number of mRNAs, \mathbf{f} is the function for modelling the interactions between \mathbf{Y} and \mathbf{X} , $\mathbf{X} \in \mathcal{R}^{N \times M}$ is the collections of the miRNA expression data where M is the number of miRNAs, and \mathbf{E} is the noise matrix. The model is very rich in presenting the miRNA targets and the function \mathbf{f} could be non-linear. However in this paper we only explore the linear model which could then be represented as

$$\mathbf{Y} = \mathbf{X}\mathbf{f} + \mathbf{E}$$

where now $\mathbf{f} \in \mathcal{R}^{M \times P}$ is a matrix. This model is similar to the bilinear model proposed in Zhong et al. (2011), where both the \mathbf{X} and \mathbf{f} could be unknown or \mathbf{X} is known and \mathbf{f} is unknown. In this paper, since both the miRNA and mRNA expression data are observed, we will consider the later case. An illustrate plot for the model is shown in the Figure 1, where the plot (A) models the direct interactions of mRNA and miRNA and (B) models the interactions of mRNA and RISC which is a complex composed of miRNA and Argonaute protein. Stanhope et al. (2009) and Lu et al. (2011) have studied the model (B) and shown the advantages of this model comparing to the model (A). We will consider the both models. So in this paper, the matrix \mathbf{X} represents the miRNA for model (A) or the complex RISC for model (B). Note that \mathbf{X} may be the multiplication of -1 and the miRNA expressions and thus a positive value of f_{mp} represents the negative regulation of miRNA. Research has found that miRNA usually takes part in negative regulation in translational repression or target degradation and gene silencing (Bartel, 2009; Kusenda et al., 2006), and however miRNA may also possibly act as positive regulation such as the transcriptional and translational activations (Vasudevan et al., 2007). So in order to model the roles of miRNA, we assume two situations $f_{mp} \geq 0$ and $f \in \mathcal{R}$.

2.1 Modelling both the positive and negative regulations

As was noted we could assume that miRNAs act as both positive and negative regulations to their targets. Then we consider one gene denoted by $\mathbf{y} \in \mathcal{R}^{N \times 1}$, and $\beta \in \mathcal{R}^{M \times 1}$ is one column of \mathbf{f} . The model is then represented as

$$\mathbf{y} = \mathbf{X}\beta + \epsilon$$

where ϵ is one column of the noise matrix \mathbf{E} . We assume that $\epsilon \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ is Gaussian noise with zero mean and variances σ^2 . The main task is then to estimate the parameter vector β which models the miRNA targets. Several algorithms are available for resolving this problem, such as the Least Square Regression, Ridge Regression, The Least Absolute Shrinkage and Selection Operator (LASSO) and Bayesian LASSO.

2.1.1 Least Square Regression

The LSR for the linear model is defined by the following optimization problem

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N \left(y_i - \sum_j \beta_j x_{ij} \right)^2 \right\}$$

The solution has the following closed form

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

2.1.2 Ridge Regression

The ridge regression for the linear model is defined by,

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N \left(y_i - \sum_j \beta_j x_{ij} \right)^2 \right\}$$

subject to

$$\sum_j \beta_j^2 \leq t$$

The solution could be represented as

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$$

where λ is a regulation parameter.

2.1.3 LASSO

The LASSO estimate for the linear regression model is defined by,

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N \left(y_i - \sum_j \beta_j x_{ij} \right)^2 \right\}$$

subject to

$$\sum_j |\beta_j| \leq t$$

We can see that no constraints are imposed on LSR, the RR uses a l_2 -norm constrain on the β and while LASSO employs the l_1 -norm. A few softwares are available for resolving LASSO and here we employed the Glmnet for Matlab (Friedman et al., 2007, 2010).

2.1.4 The Bayesian LASSO

While LSR, RR and LASSO are point estimate algorithms, we also examine the Bayesian LASSO approach of Park and Casella (2008) which infers the posterior densities for the parameters β . Given the linear regression model, the data likelihood follows a Normal distribution with the following form

$$p(\mathbf{y}|\mathbf{X}, \beta, \sigma^2) = \mathcal{N}(\mathbf{X}\beta, \sigma^2 \mathbf{I}_N)$$

We employ the following priors for those parameters

$$p(\beta|\sigma^2, \tau_1^2, \dots, \tau_M^2) = \mathcal{N}(\mathbf{0}_M, \sigma^2 D_\tau),$$

where $D_\tau = \text{diag}(\tau_1^2, \dots, \tau_M^2)$

$$p(\sigma^2) \sim \frac{1}{\sigma^2}$$

$$p(\tau^2|\lambda) = \prod_{j=1}^M \frac{\lambda^2}{2} \exp\left(-\frac{1}{2}\lambda^2 \tau_j^2\right) d\tau_j^2$$

$$p(\lambda) \propto (\lambda^2)^{r-1} e^{-\frac{\lambda^2}{\delta}}$$

So given the likelihood and priors, we are interested in simulating the conditional distribution

$$p(\beta, \sigma^{-2}, \tau_j^{-2}, \lambda^2 | \mathbf{y}, \mathbf{X}) \propto p(\mathbf{y}|\mathbf{X}, \beta, \sigma^2) p(\beta|\sigma^2, \tau^2) p(\tau^2|\lambda) p(\sigma^2) p(\lambda)$$

The posterior distributions are then

$$p(\beta|\mathbf{X}, \mathbf{y}) = \mathcal{N}(\mu_\beta, \Sigma_\beta)$$

where $\mu_\beta = A^{-1} \mathbf{X}^T \mathbf{y}$, $\Sigma_\beta = \sigma^2 A^{-1}$, and $A = \mathbf{X}^T \mathbf{X} + D_\tau^{-1}$.

$$p(\sigma^{-2} | \mathbf{X}, \mathbf{y}) \sim \text{Gamma}(\alpha_\sigma, \beta_\sigma)$$

where $\alpha_\sigma = \frac{1}{2}(N+M)+2$ and $\beta_\sigma = [\frac{1}{2}\beta^T D_\tau^{-1} \beta + \frac{1}{2}(\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)]^{-1}$, and *Gamma* represents a Gamma distribution with the density $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$.

$$p(\tau_j^{-2} | \dots) \sim \text{InvGaussian}(\mu_\tau, \lambda_\tau)$$

where $\mu_\tau = \sqrt{\frac{\lambda^2 \sigma^2}{\beta_j^2}}$ and $\lambda_\tau = \lambda^2$, and *InvGaussian* represents an inverse Gaussian distribution which has a density given by $f(x) = \sqrt{\frac{\lambda}{2\pi}} x^{-\frac{3}{2}} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}$ where $x > 0$.

$$p(\lambda^2 | \dots) \sim \text{Gamma}(\alpha_\lambda, \beta_\lambda)$$

where $\alpha_\lambda = M+r$ and $\beta_\lambda = \left(\frac{1}{\delta} + \frac{1}{2} \sum_{j=1}^M \tau_j^2\right)^{-1}$. It can be seen that all the posteriors are the known distributions and thus we could simulate the parameters by using the Gibbs sampler in the Algorithm 1.

Algorithm 1 Gibbs Sampler for Bayesian LASSO

Require: \mathbf{y} , \mathbf{X} , $r = 1$, $\delta = 1$, NSamps

for $t=1:\text{NSamps}$ **do**

 Draw sample $\tau_j^{-2} \sim \text{InvGaussian}(\mu_\tau, \lambda_\tau)$ for all j

 Draw sample $\beta \sim \mathcal{N}(\mu_\beta, \Sigma_\beta)$

 Draw sample $\sigma^{-2} \sim \text{Gamma}(\alpha_\sigma, \beta_\sigma)$

 Draw sample $\lambda^2 \sim \text{Gamma}(\alpha_\lambda, \beta_\lambda)$

end for

2.2 Modelling the negative regulations

The negative interactions of miRNA and mRNA are modeled by the following linear regression model when β is non-negative,

$$\mathbf{y} = -\mathbf{X}\beta + \epsilon$$

where now \mathbf{X} is the miRNA expression data. To resolve this model, we employ both the point estimate algorithm and the Bayesian solution.

2.2.1 Non-negative LASSO Algorithm

The nLASSO for the linear regression model is defined by

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N \left(y_i + \sum_j \beta_j x_{ij} \right)^2 \right\}$$

subject to

$$\sum_j \beta_j \leq t, \beta_j \geq 0$$

We see that the only difference between LASSO and nLASSO is that nLASSO also employs the non-negative constraints. For implementing this algorithm, we employ the Matlab software available at http://www.stanford.edu/~boyd/l1_ls/ (Kim et al., 2007).

2.2.2 Non-negative Bayesian LASSO

Similar to the Bayesian LASSO, we could also simulate the negative regulators by using sampling approaches. To simulate the required posterior distribution, we employ the following data likelihood and the priors for those parameters

$$p(\mathbf{y}|\mathbf{X}, \beta, \sigma^2) = \mathcal{N}(-\mathbf{X}\beta, \sigma^2 \mathbf{I})$$

$$p(\beta|\sigma^2, \lambda) = \prod_{m=1}^M \frac{\lambda_m}{2\sigma^2} \exp \left\{ -\frac{\lambda_m \beta_m}{2\sigma^2} \right\} \mathbf{I}(\beta_m \geq 0)$$

$$p(\lambda_m) = \frac{1}{\Gamma(\alpha_\lambda^0)(\beta_\lambda^0)^{\alpha_\lambda^0}} \lambda_m^{\alpha_\lambda^0-1} e^{-\frac{\lambda_m}{\beta_\lambda^0}} \mathbf{I}(\lambda_m > 0)$$

$$p(\sigma^2) \propto \frac{1}{\sigma^2}$$

We are interested in simulating the conditional distribution

$$p(\beta, \sigma^2, \lambda | \mathbf{y}, \mathbf{X}) \propto p(\mathbf{y} | \mathbf{X}, \beta, \sigma^2) p(\beta | \sigma^2, \lambda) p(\sigma^2) p(\lambda)$$

The conditional distribution for β is then a truncated Normal

$$p(\beta | \mathbf{y}, \mathbf{X}) \propto \mathcal{N}_+(\mu_\beta, \Sigma_\beta), \beta_i \geq 0$$

where $\mu_\beta = -\Sigma_X(\mathbf{X}^T \mathbf{y} + \frac{1}{2} \boldsymbol{\lambda})$, $\Sigma_\beta = \sigma^2 \Sigma_X$ where $\Sigma_X = (\mathbf{X}^T \mathbf{X})^{-1}$, and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_M)^T$. Sampling from the truncated Normal distribution is shown in the Appendix. The conditional distribution for $\sigma^{-2} \sim \text{Gamma}(\alpha_\sigma, \beta_\sigma)$ is a Gamma distribution where $\beta_\sigma = \left[\frac{1}{2} (\tilde{\mathbf{y}} + \mathbf{X}\beta)^T (\tilde{\mathbf{y}} + \mathbf{X}\beta) + \frac{1}{2} \sum_{m=1}^M \lambda_m \beta_m \right]^{-1}$ and $\alpha_\sigma = \frac{1}{2} N + M + 2$. The conditional for λ_m is also a Gamma distribution, i.e. $\lambda_m \sim \text{Gamma}(\alpha_\lambda, \beta_\lambda)$, where $\alpha_\lambda = \alpha_\lambda^0 + 1$ and $\beta_\lambda = \left(\frac{1}{\beta_\lambda^0} + \frac{1}{2\sigma^2} \beta_m \right)^{-1}$. So finally we use the Gibbs sampler in Algorithm 2 for simulating those parameters.

Algorithm 2 Gibbs Sampler for Non-negative Bayesian LASSO

Require: $\mathbf{y}, \mathbf{X}, \alpha_\lambda^0 = 1e-6, \beta_\lambda^0 = 1e6, \text{NSamps}$
for $t=1:\text{NSamps}$ **do**
 Draw sample $\beta \sim \mathcal{N}_+(\mu_\beta, \Sigma_\beta)$
 Draw sample $\sigma^{-2} \sim \text{Gamma}(\alpha_\sigma, \beta_\sigma)$
 Draw sample $\lambda_m \sim \text{Gamma}(\alpha_\lambda, \beta_\lambda)$ for $m = 1, 2, \dots, M$.
end for

3 A Variable Selection Scheme

Both the Bayesian Lasso and the non-negative Bayesian Lasso are not able to select variables automatically. Here we propose a particular scheme for variable selection for this problem. In Bayesian statistics, credible interval is usually used to summary the statistics for the inferred posterior distribution of a parameter (See for example Eberly and Casella (2003)). We define a credible interval $[a, b]$ for a variable. If a is significantly greater than zero, we assume that the variable is active (or selected). So we call it the active credible interval (ACI). Since many of the posterior distributions are multimodal, we partition the posterior samples into two clusters. We observed that for this particular problem the posterior of β has high probability density around zero and also it may have high probability density apart from zero. So if the posterior really has high probability density greater than zero, we would assume the corresponding variable is active. Ideally, the samples of the posterior are partitioned into two clusters, where one cluster

has high probability density around zero and the other is apart from zero. We calculated the ACI. The statistical significance of the ACI is then defined as

$$p = \frac{Q}{T}$$

where Q is the number of samples in the ACI and T is the totally number of the samples. For fixing a particular value $0 < \alpha < 1$, if $p > \alpha$, we then assume that the variable is selected (or active). Obviously, the value of α controls the numbers of active variables. As the value of α increasing, the smaller numbers of variables are selected.

4 Expression Data Sets

The experimentally validated targets of the four data sets were employed to validate those linear regression algorithms. The first data, which is noted as Multi Class Cancer (MCC), has been studied by Muniategui et al. (2012a). In this data, 23 mRNA-miRNA interactions have been identified by searching from the literatures. This data set has 88 paired cancer and normal samples with mRNA data from Ramaswamy et al. (2001) and miRNA expression data from Lu et al. (2005). We employed the experimentally validated targets reported in Muniategui et al. (2012a) and the method there for selecting the initial set of putative interactions. The following three data sets, which are also using Argonaute protein expression, were studied by Lu et al. (2011) and Stanhope et al. (2009). The second data was to study the nasopharyngeal cancer by those researchers in Madison, WI, USA which derived mRNA and miRNA profiling from 31 NPC and 10 normal tissue samples. The third data set was from the Broad institute which contains 67 tissue samples of 10 different normal and tumor tissue types. The fourth data set was conducted by Memorial Sloan-Kettering Cancer Center (MSKCC) with 28 normal, 98 primary cancer and 13 metastatic cancer tissues samples using the Affymetrix Human Exon 1.0 ST Array and Agilent Human miRNA Microarray 2.0. We use the experimentally detected mRNA-miRNA interactions for the data sets Broad, Madison and MSKCC collected by (Lu et al., 2011) to compare those algorithms. The numbers of experimentally validated interactions are listed in the Table 7 for all the data sets.

5 Results

In this section we apply six linear regression algorithms to four expression data sets. The six linear regression algorithms are least square regression, ridge regression, LASSO, non-negative LASSO, Bayesian LASSO and non-negative Bayesian LASSO. The first four algorithms produce point estimates for the model and the last two methods are Bayesian methods.

5.1 Bayesian Methods Perform Better in Modelling Sparse Interactions for mRNA and miRNA

We have applied the six linear regression algorithms to the known targets of the four data sets described. The four algorithms, i.e. LSR, ridge regression, LASSO and non-negative LASSO, produce point estimates for those parameters. Across all the data sets employed, the LSR, ridge regression and LASSO produced the same results. We observed that across all the data sets, LSR, ridge regression and LASSO algorithms were not able to learn sparsity for the model (see for example Figures 2 & 4) and thus additional schemes such as the one proposed in Lu et al. (2011) have to be employed for the variable selection purposes. As was reported in Muniategui et al. (2012a), we also confirm that the nLASSO was able to learn model sparsity. See the Figures 2 & 4. Figure 2 shows the learned β for the gene TWIST1 where there are 21 candidate microRNAs. Looking at this example, the non-negative LASSO algorithm clearly identified three miRNAs and they are miR-137, miR-145 and miR-200c and all the coefficients for the rest miRNAs tend to zero. However, the LSR, Ridge regression and LASSO estimated 10 negative values for β and 11 positive values and all those values were not tending to zero. Thus, 11 miRNAs were identified for down regulating TWIST1. This shows that the algorithms with the non-negative constraint on the parameters is more suitable for modelling sparse interactions of mRNA and miRNA than those without any constraints, since we are mainly looking for miRNAs which down regulate mRNAs. Although LASSO algorithm is not considering model sparsity for this particular problem, it identified more miRNAs than nLASSO. This may be an advantage for employing LASSO algorithm, comparing to the nLASSO algorithm, since it selects more miRNAs. Comparing to those point estimate algorithms, we can infer the posterior densities for those parameters for the gene TWIST1 using the BLASSO and nBLASSO. Comparing to the LASSO algorithm, however, the Bayesian LASSO where no non-negative constraints were employed was also able to produce convincing results for identifying mRNA-miRNA interactions. This may be due to that Bayesian LASSO was able to explore the whole parameter spaces, and while the LASSO is required to tune the regularizing parameter λ which is crucial for learning the model sparsity. The non-negative Bayesian LASSO performs best among those algorithms employed. Figure 3 shows the inferred densities for five miRNAs. Looking at those densities, for the ones inferred by BLASSO, we would confidently select miR-145, miR-200b and miR-200c and ignore miR-141 and miR-200a since the densities of miR-141 and miR-200a seem symmetric around zero. While for the ones inferred by nBLASSO, we would confidently select miR-200b and miR-200c since the densities for them have two peaks at around both zero and 0.05 and we are not sure to select miR-141, miR-145 and miR-200a since they have only one peak at around zero. We then need to calculate the active credible intervals and their statistical significance to perform the variable selections. For the gene TWIST1, the results for the active credible intervals and their statistical significance are shown in the Table 1 for the Bayesian LASSO and the Table 2 for the non-negative Bayesian

LASSO. If we set the parameter $\alpha = 0.05$, the BLASSO selected miR-145, miR-200b and miR-200c and while the nBLASSO selected miR-141, miR-145, miR-200a, miR-200b and miR-200c. It should be noted that the miR-200 family (miR-200a, miR-200b, miR-200c and miR-141) were experimentally identified as down-regulating TWIST1 (Wiklund et al., 2011). We analyzed all the genes using the same procedure.

We show all the credible intervals and their statistical significance which are greater than 0.05 for the experimentally validated genes of the MCC data set computed by BLASSO and nBLASSO in the Tables 6 & 5, respectively. We set $\alpha = 0.05$ for selecting the miRNAs. Note that to perform variable selection, we have to manually fix a threshold for the point estimate algorithms such as the LASSO and nLASSO. We see that nBLASSO selected more miRNAs than BLASSO, which is due to the non-negative constraints. We also counted the numbers of experimentally validated interactions estimated by the LASSO, nLASSO, BLASSO and nBLASSO for all the four data sets which are shown in the Table 7. Note that the analysis for the three data sets Broad, Madison and MSKCC which are using Argonaute protein information is shown in the following sections. We see that the Bayesian approaches identified more miRNA and mRNA interactions than the point estimate algorithms across all the data sets, which conclude that the Bayesian approaches perform better than the point estimate algorithms in modelling the mRNA and miRNA interactions.

5.2 Bayesian Analysis for mRNA-miRNA Interactions Using Argonaute Protein Information

We have shown that both the Bayesian approaches and the non-negative LASSO for linear regression model are able to identify the sparse representations for the miRNA-mRNA interactions. It has been shown that the regression models employing Argonaute expressions fit the expression profiles of known targets pairs substantially better than models based only on miRNA expression data (Stanhope et al., 2009; Lu et al., 2011). We applied all the six algorithms to three expression data sets, which are Broad, Madison and MSKCC, involving Argonaute expression information, which were also studied in Stanhope et al. (2009) and Lu et al. (2011). Here we analyze the results for the gene NOTCH1 and we analyze all the rest genes using the same method. Twelve miRNAs were selected as the candidates for interacting with the gene NOTCH1. Six miRNAs have been experimentally detected to down regulate NOTCH1 and they are miRNA 34a, 34b & 34c, miRNA 23b, miRNA 24 and miRNA 27b (Fukuda et al., 2005; Wang et al., 2010; Zheng et al., 2012; Melo and Kalluri, 2012). The point estimate algorithms, i.e., the LSR, ridge regression, LASSO and non-negative LASSO, were applied to the expression data for identifying the interactions of the mRNA and miRNA. The estimated β are shown in the Figure 4. In the figure, the black bars plot the β for the miRNA involving Argonaute 2 expression information and the red bars plot the β for the miRNA

involving Argonaute 1,3&4 expression information. Obviously the LSR, ridge regression and LASSO produce similar results. Since the Argonaute 2 protein is the competitor of Argonaute 1,3&4 in forming the RISC for targeting mRNA, as was discussed in Stanhope et al. (2009) the values of β for miRNA with Ago2 information computed by LSR, ridge regression and LASSO are inversely related to that of miRNA with Ago1,3,&4 information in sign with similar magnitudes of the β values. For the results of the non-negative LASSO, if the effect of a miRNA with Ago2 is positive, the effect of the miRNA with Ago1,3&4 then is zero, and vice versa, and also both the effects may be zero. For the LSR, ridge regression and LASSO, we would prefer selecting all of the candidate miRNAs, and while the non-negative LASSO selects smaller number of miRNAs. We also applied the Bayesian approaches to this data. The probability densities are plotted in the Figures 5 & 6 for Bayesian LASSO and non-negative Bayesian LASSO, respectively. We then calculated the active credible intervals and as well as the statistical significance which are shown in the Tables 3 & 4. We use the parameter $\alpha = 0.05$. The Bayesian LASSO selected seven miRNAs combined with Ago2 and five miRNAs combined with Ago1,3&4 which contains 5 known targets out of 6 and the nBLASSO selected eight miRNAs combined with Ago2 and seven miRNAs combined with Ago1,3&4 which contains 5 out of 6 known targets. Their densities are shown in Figures 5 & 6. For the purpose of clearly comparing the effects of Ago2 and Ago1,3&4, we show the results for gene ACAA2 where only the miRNA-124 was selected as the candidate for targeting ACAA2. Figure 7 plots the traces for sampling the effects of miRNA-124 \times Ago2 and miRNA-124 \times Ago1,3&4. Interestingly, when the effects of miRNA-124 \times Ago2 are on, the effects of miRNA-124 \times Ago1,3&4 are then off and vice versa. This implies that Ago2 and Ago1,3&4 are competing with each other to form the RISC which agrees the results of Stanhope et al. (2009).

6 Conclusions

As linear regression models have been proved suitable for modelling miRNA targets, we have proposed Bayesian LASSO and non-negative Bayesian LASSO for inferring mRNA-miRNA interactions. We have applied the least square regression, ridge regression, LASSO, non-negative LASSO, Bayesian LASSO and non-negative Bayesian LASSO to four publicly available expression data sets. We conclude that LSR, RR and LASSO algorithms perform the same, and the nLASSO produces sparse representations for the mRNA-miRNA interactions since the non-negative constraints were imposed. However we argue that LSR, RR and LASSO estimate more numbers of mRNA-miRNA interactions than nLASSO. Both the BLASSO and nBLASSO provide uncertainties for the estimates of the interactions. The Bayesian methods do not perform variable selections automatically and thus we proposed to employ the credible intervals to accomplish this task by looking at a statistical significance. Of course, the point estimate algorithms are also lack of variable selection automatically and thus a threshold may have to be manually fixed for selecting miRNAs. However, we

could employ advanced methodologies for performing variable selections for this problem as was proposed in Zhong et al. (2011). We prefer the Bayesian approaches rather than the point estimate algorithms, because Bayesian methods produce the probability densities for the effects of the miRNAs to their targets. So we can analyze the mRNA-miRNA interactions by graphing those inferred densities of the effects. It has been well known that the Ago2 proteins are competitors of the Ago1,3&4 proteins in binding with RISC. Our Gibbs sampler was able to simulate this competition by sampling the effects of the RISC to miRNA targets. Overall, comparing to point estimate algorithms, Bayesian methods were able to infer more numbers of known mRNA and miRNA interactions.

Appendix

In order to simulate the truncated Normal distribution, we employ a Gibbs sampler. We consider sampling from the general truncated Normal distribution of the form

$$p(\mathbf{z}|\mu, \Sigma) \propto \exp\left(-\frac{1}{2}(\mathbf{z} - \mu)^T \Sigma^{-1}(\mathbf{z} - \mu)\right) \mathbf{I}(\mathbf{z} \in C)$$

We are to sample z_i given $z_{-i} = (z_1, z_2, \dots, z_{i-1}, z_{i+1}, \dots, z_M)$. The conditional distribution is then

$$p(z_i|z_{-i}) \propto \exp\left(-\frac{1}{2v_i^2}(z_i - u_i)^2\right) I(z_i \in (c_i, d_i))$$

where $v_i^2 = \omega_{ii}^{-1}$, $u_i = \mu_i + \frac{1}{\omega_{ii}} \sum_{j \neq i} (\mu_j - z_j) \omega_{ij}$ where ω_{ij} is the ij th element of Σ^{-1} . Set $\xi = \frac{z_i - u_i}{v_i}$, it is required to simulate $p(\xi) \propto \exp(-\xi^2/2) I(\xi \in (c_\xi, d_\xi))$ where $c_\xi = \frac{c_i - u_i}{v_i}$ and $d_\xi = \frac{d_i - u_i}{v_i}$. We employ the Gibbs sampler in Damien and Walker (2001) to simulate $p(\xi)$. Then ξ is simulated in turn according to the conditionals which are Uniform:

$$p(Y|\xi) \sim U(0, \exp(-\xi^2/2))$$

$$p(\xi|y) \sim U\left(\max\left(c_\xi, -\sqrt{-2 \log y}\right), \min\left(d_\xi, \sqrt{-2 \log y}\right)\right)$$

Acknowledgements

Mingjun Zhong is supported by the fundamental research funds for the central universities in China and the National Natural Science Foundation of China grant number 60501021.

References

DP. Bartel. MicroRNAs: Target recognition and regulatory functions. *Cell*, 136 (2):215–233, 2009.

- D Betel, M. Wilson, A. Gabow, and et al. The microRNA.org resource: targets and expression. *Nucleic Acids Research*, 36:D149–D153, 2008.
- R.W. Carthew and E.J. Sontheimer. Origins and mechanisms of miRNAs and siRNAs. *Cell*, 136:642–655, 2009.
- P. Damien and S.G. Walker. Sampling truncated Normal, Beta, and Gamma densities. *Journal of Computational and Graphical Statistics*, 10(2):206–215, 2001.
- LE. Eberly and G. Casella. Estimating Bayesian credible intervals. *Journal of Statistical Planning and Inference*, 112:115–132, 2003.
- J. Friedman, T. Hastie, H. Hofling, and R. Tibshirani. Pathwise coordinate optimization. *Annals of Applied Statistics*, 1(2):302–332, 2007.
- J. Friedman, T. Hastie, and R. Tibshirani. Regularized paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1): 1–22, 2010.
- Y Fukuda, H. Kawasaki, and K Taira. Exploration of human miRNA target genes in neuronal differentiation. *Nucleic Acids Symp Ser (Oxf)*, 38:341C342, 2005.
- S. Griffiths-Jones, HK. Saini, S. Van Dongen, and AJ. Enright. miRBase: tools for microRNA genomics. *Nucleic Acids Research*, 36:D154–158, 2008.
- JC. Huang, QD. Morris, and BJ. Frey. Bayesian inference of microRNA targets from sequence and expression data. *Journal of Computational Biology*, 14(5): 550–563, 2007.
- G. Hutvagner. Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. *FEBS Letters*, 579:5850–5857, 2005.
- S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale l1-regularized least squares. *IEEE Journal on Selected Topics in Signal Processing*, 1(4):606–617, 2007.
- A. Kozomara and S. Griffiths-Jones. miRBase: intergrating microRNA annotation and deep-sequencing data. *Nucleic Acids Reseach*, 39:D152–D157, 2011.
- A. Krek, D. Grun, MN. Poy, and et al. Combining microRNA terget predictions. *Nature Genetics*, 37:495–500, 2005.
- B. Kusenda, M. Mraz, J. Mayer, and S. Pospisilova. MicroRNA biogenesis, functionality and cancer relevance. *Biomedical Papers*, 150(2):205–215, 2006.
- Y. Lee, C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Radmark, S. Kim, and VN. Kim. The nuclear RNase III Drosha initiates microRNA processing. *Nature*, 425:415–419, 2003.

- B.P. Lewis, I.H. Shih, M.W. Jones-Rhoades, D.P. Bartel, and CB. Burge. Prediction of mammalian microRNA targets. *Cell*, 115:787–798, 2003.
- BP. Lewis, CB. Burge, and DP. Bartel. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120(1):15–20, 2005.
- J. Lu, G. Getz, EA Miska, E. Alvarez-Saavedra, J. Lamb, and et al. MicroRNA expression profiles classify human cancers. *Nature*, 435(7043):834–838, 2005.
- Y. Lu, Y. Zhou, Qu W., Deng M., and Zhang C. A lasso regression models for the construction of microRNA-target regularoty networks. *Bioinformatics*, 27(17):2406–2413, 2011.
- SA. Melo and R. Kalluri. Angiogenesis is controlled by miR-27b associated with endothelial tip cells. *Blood*, 119(11):2439–2440, 2012.
- A. Muniategui, R. Nogales-Cadenas, M. Vázquez, X. L. Aranguren, X. Agirre, A. Luttun, F. Prosper, A. Pascual-Montano, and A. Rubio. Quantification of miRNA-mRNA interactions. *PLOS one*, 7(2):e30766, 2012a.
- A. Muniategui, J. Pey, F. Planes, and A. Rubio. Joint analysis of miRNA and mRNA expression data. *Briefings in Bioinformatics*, pages 1–16, 2012b.
- T. Park and G. Casella. The Bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008.
- R. Parker and U. Sheth. P bodies and the control of mRNA translation and degradation. *Molecular Cell*, 25:635–646, 2007.
- R.S. Pillai, S.N. Bhattacharyya, and W. Filipowicz. Repression of protein synthesis by miRNAs: how many mechanisms? *Trends in Cell Biology*, 17: 118–126, 2007.
- S. Ramaswamy, P. Tamayo, R. Rifkin, S. Mukherjee, and CH. Yeang. Multiclass cancer diagnosis using tumor gene expression signatures. *Proc Natl Acad Sci USA*, 98(26):15149–15154, 2001.
- SA. Stanhope, S. Sengupta, J. den Boon, P. Ahlquist, and MA. Newton. Statistical use of argonaute expression and RISC assembly in microRNA target identification. *PLOS Computational Biology*, 5(9):e1000516, 2009.
- F. Stingo, YA. Chen, M. Vannucci, M. Barrier, and PE. Mirkes. A Bayesian graphical modeling approach to microRNA regulatory network inference. *The Annals of Applied Statistics*, 4(4):2024–2048, 2010.
- N. Su, Y. Wang, M. Qian, and M. Deng. Predicting microRNA targets by integrating sequence and expression data in cancer. *IEEE Int Conf Syst Biol*, 2011.

- R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, 58(1):267–288, 1996.
- S. Vasudevan, Y. Tong, and JA. Steitz. Switching from repression to activation: microRNAs can up-regulate translation. *Science*, 318:1931–1934, 2007.
- Z. Wang, Y. Li, D. Kong, A. Ahmad, S. Banerjee, and FH. Sarkar. Cross-talk between miRNA and Notch signaling pathways in tumor development and progression. *Cancer Letters*, 292:141–148, 2010.
- ED. Wiklund, JB. Bramsen, T. Hulf, L. Dyrskjot, R. Ramanathan, TB. Hansen, SB. Villadsen, S. Gao, MS. Ostenfeld, M. Borre, ME. Peter, TF. orntoft, J. Kjems, and SJ. Clark. Coordinated epigenetic repression of the mir-200 family and mir-205 in invasive bladder cancer. *International Journal of Cancer*, 128(6):1327–1334, 2011.
- J. Zheng, H.-Y. Jiang, J. Li, H.-C. Tang, X.-M. Zhang, X.-R. Wang, J.-T. Du, H.-B. Li, and G. Xu. MicroRNA-23b promotes tolerogenic properties of dendritic cells in vitro through inhibiting Notch1/NF-B signalling pathways. *European Journal of Allergy and Clinical Immunology*, 67:362–370, 2012.
- M. Zhong, M. Girolami, K. Faulds, and D. Graham. Bayesian methods to detect dye-labelled DNA oligonucleotides in multiplexed raman spectra. *Journal of the Royal Statistical Society Series C*, 60(2):187–206, 2011.

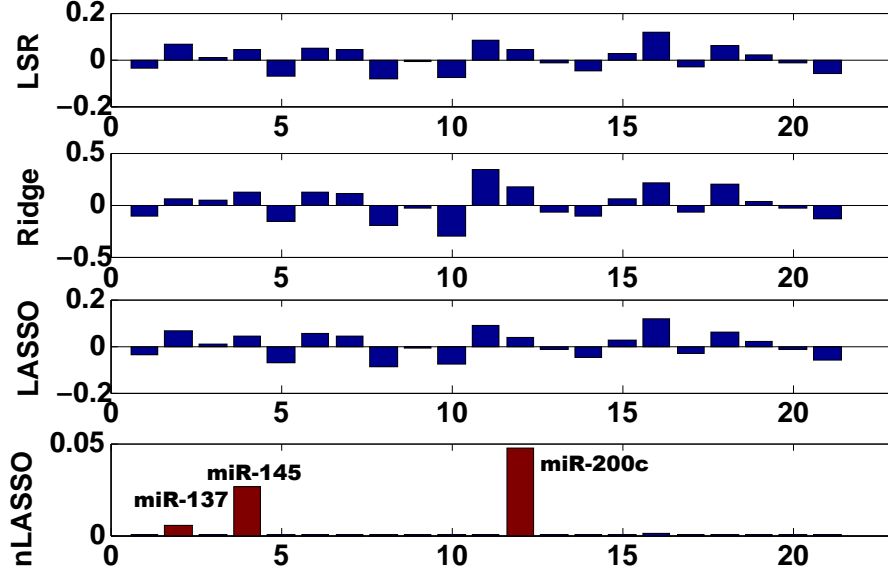


Figure 2: The point estimates for β for the gene TWIST1 in the MCC data set using Least Square Regression, Ridge Regression, LASSO and Non-negative LASSO algorithms.

mRNA	miRNA	Significance	Credible Interval
TWIST1	miR-145	0.104	[0.018,0.068]
	miR-200b	0.153	[0.028,0.083]
	miR-200c	0.691	[0.032,0.086]

Table 1: The active credible intervals and their statistical significance computed by Bayesian LASSO for gene TWIST1 ($\alpha = 0.05$). The miRNAs highlighted by red were experimentally validated targeting TWIST1.

mRNA	miRNA	Significance	Credible Interval
TWIST1	miR-141	0.073	[0.023,0.113]
	miR-145	0.104	[0.026,0.132]
	miR-200a	0.058	[0.023,0.118]
	miR-200b	0.193	[0.028,0.111]
	miR-200c	0.390	[0.028,0.096]

Table 2: The active credible intervals and their statistical significance computed by non-negative Bayesian LASSO for gene TWIST1 ($\alpha = 0.05$). The miRNAs highlighted by red were experimentally validated targeting TWIST1.

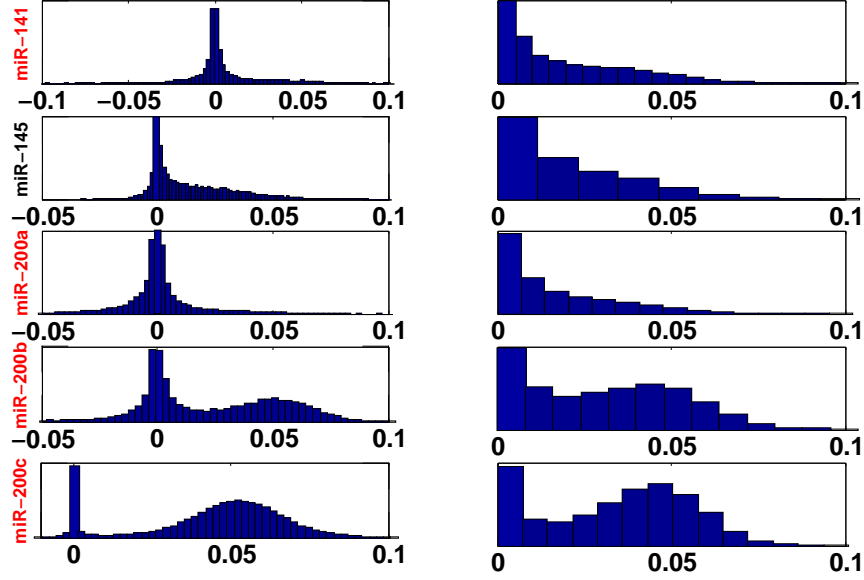


Figure 3: The inferred probability densities for β by using the Bayesian LASSO (left) and non-negative Bayesian LASSO (right). The miRNAs highlighted by red were experimentally validated as down regulating the gene TWIST1.

mRNA	miRNA	Significance	Credible Interval
NOTCH1	miR-92 \times Ago2	0.058	[0.048,0.217]
	miR-139 \times Ago2	0.056	[0.054,0.171]
	miR-30d \times Ago2	0.162	[0.071,0.198]
	miR-34c \times Ago2	0.262	[0.041,0.138]
	miR-23b \times Ago2	0.182	[0.064,0.203]
	miR-24 \times Ago2	0.234	[0.067,0.220]
	miR-27b \times Ago2	0.066	[0.039,0.154]
	miR-92 \times Ago134	0.078	[0.024,0.086]
	miR-139 \times Ago134	0.154	[0.046,0.167]
	miR-30d \times Ago134	0.530	[0.071,0.191]
	miR-34b \times Ago134	0.118	[0.047,0.146]
	miR-24 \times Ago134	0.138	[0.051,0.207]

Table 3: The inferred active credible intervals and their statistical significance computed by Bayesian LASSO ($\alpha = 0.05$). The miRNAs highlighted by red were experimentally validated to down regulate NOTCH1.

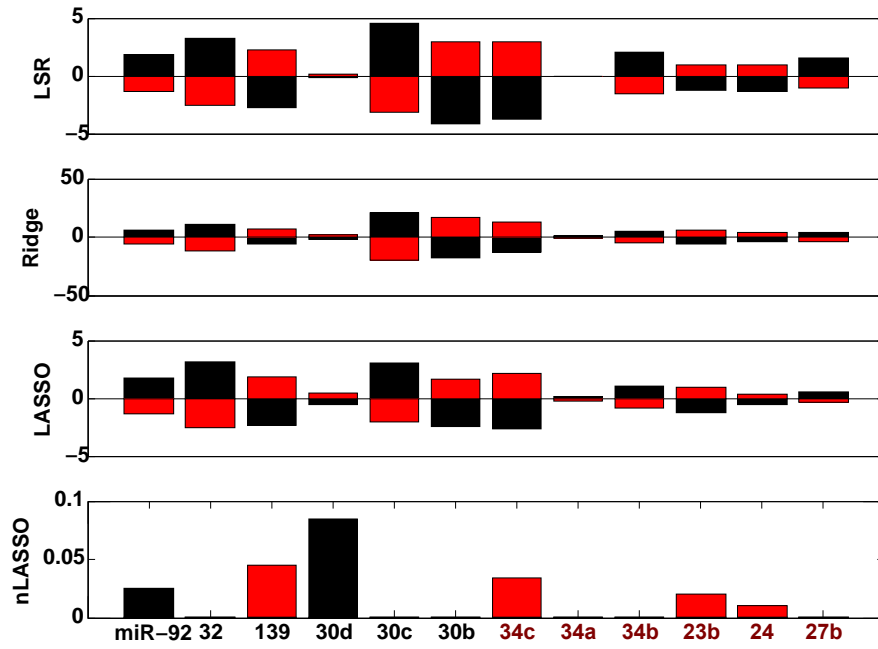


Figure 4: The point estimates for β using Least Square Regression, Ridge Regression, LASSO and Non-negative LASSO algorithms for gene NOTCH1 using Argonaute expression, where black bar is the values for miRNA using Argonaute 2 and while the red is using Argonaute 1, 3, and 4.

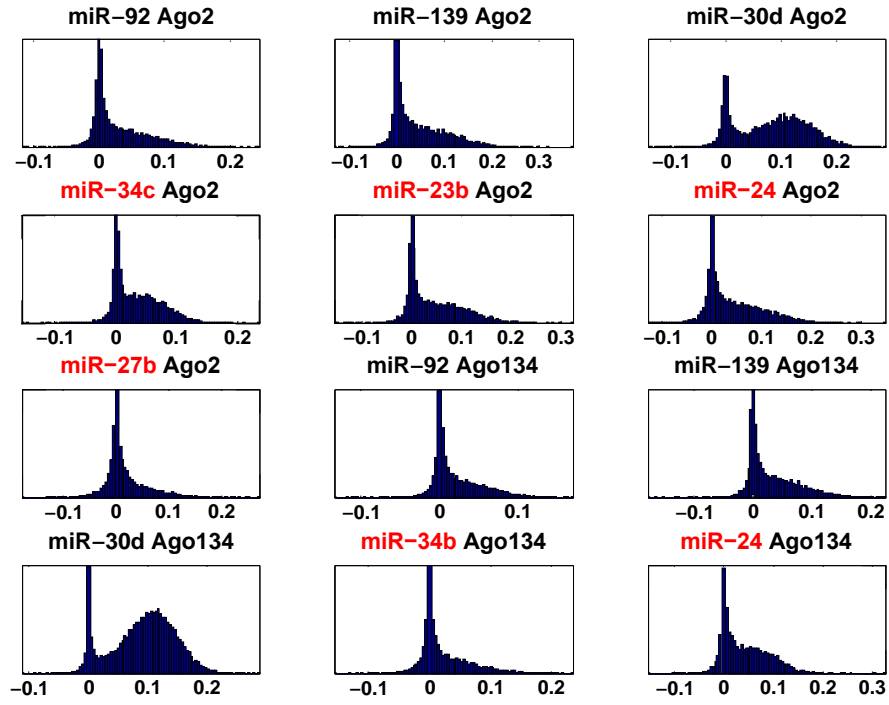
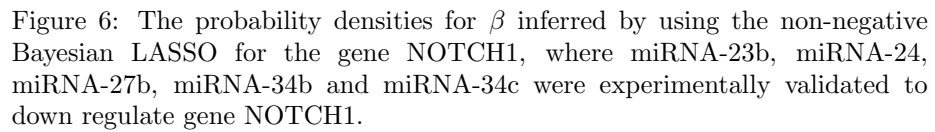


Figure 5: The probability densities for β inferred by using the Bayesian LASSO for the gene NOTCH1, where miRNA-23b, miRNA-24, miRNA-27b, miRNA-34b and miRNA-34c were experimentally validated to down regulate gene NOTCH1.



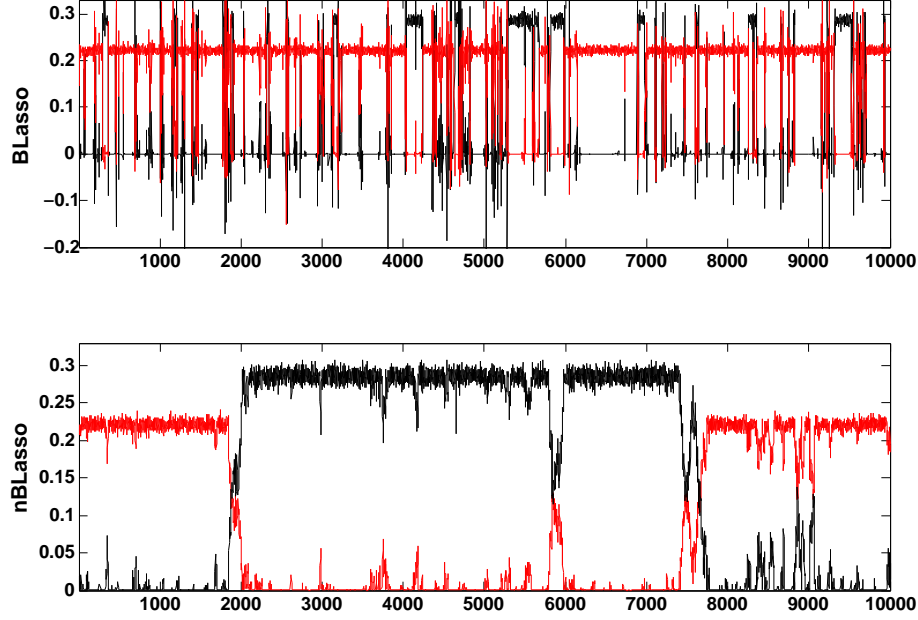


Figure 7: The samples for β drawn by using Bayesian LASSO and non-negative Bayesian LASSO, where the black plots the β for miRNA-124 \times Ago2 and the red plots for miRNA-124 \times Ago134.

mRNA	miRNA	Significance	Credible Interval
NOTCH1	miR-92 \times Ago2	0.124	[0.046,0.156]
	miR-139 \times Ago2	0.175	[0.061,0.203]
	miR-30d \times Ago2	0.249	[0.059,0.177]
	miR-34c \times Ago2	0.102	[0.033,0.118]
	miR-34b \times Ago2	0.054	[0.043,0.158]
	miR-23b \times Ago2	0.098	[0.036,0.136]
	miR-24 \times Ago2	0.191	[0.051,0.168]
	miR-27b \times Ago2	0.051	[0.022,0.077]
	miR-92 \times Ago134	0.093	[0.028,0.109]
	miR-139 \times Ago134	0.121	[0.042,0.146]
	miR-30d \times Ago134	0.262	[0.047,0.145]
	miR-34c \times Ago134	0.121	[0.026,0.097]
	miR-34b \times Ago134	0.071	[0.032,0.137]
	miR-23b \times Ago134	0.081	[0.021,0.083]
	miR-24 \times Ago134	0.143	[0.038,0.117]

Table 4: The inferred active credible intervals and their statistical significance computed by non-negative Bayesian LASSO ($\alpha = 0.05$). The miRNAs highlighted by red were experimentally validated to down regulate NOTCH1.

mRNA	miRNA	Significance	Credible Interval	mRNA	miRNA	Significance	Credible Interval
EEF1A2	let-7f	0.628	[0.107, 0.239]	CDKN2A	miR-125b	0.155	[0.051, 0.197]
	miR-181b	0.067	[0.046, 0.178]		miR-195	0.057	[0.033, 0.151]
	miR-182	0.066	[0.037, 0.151]		miR-29a	0.051	[0.066, 0.305]
FSCN1	miR-133a	0.458	[0.071, 0.201]	CEACAM5	miR-99a	0.085	[0.032, 0.135]
	miR-145	0.103	[0.041, 0.148]		mir-143	0.079	[0.058, 0.223]
	miR-29c	0.291	[0.084, 0.262]		miR-145	0.096	[0.068, 0.254]
BRAF	let-7f	0.053	[0.037, 0.154]	KRT14	miR-221	0.057	[0.071, 0.283]
	miR-145	0.150	[0.041, 0.143]		miR-143	0.054	[0.029, 0.123]
	miR-183	0.059	[0.038, 0.157]	IGFBP6	miR-181b	0.115	[0.051, 0.182]
	miR-192	0.084	[0.016, 0.063]		miR-27a	0.073	[0.048, 0.185]
	miR-221	0.103	[0.049, 0.188]	PIGR	miR-125b	0.279	[0.048, 0.152]
	miR-223	0.318	[0.063, 0.200]		let-7b	0.314	[0.125, 0.399]
	miR-9	0.107	[0.044, 0.167]	PLK1	miR-100	0.411	[0.046, 0.136]
CCND1	miR-130b	0.171	[0.176, 0.702]		miR-99a	0.129	[0.031, 0.107]
	miR-194	0.097	[0.031, 0.138]	BAK1	miR-99b	0.164	[0.075, 0.247]
	miR-7	0.129	[0.061, 0.270]		miR-125b	0.431	[0.038, 0.109]
HOXD10	let-7b	0.101	[0.124, 0.595]	E2F1	miR-16	0.308	[0.053, 0.164]
	miR-103	0.050	[0.056, 0.277]		miR-25	0.117	[0.038, 0.137]
	miR-125b	0.088	[0.041, 0.203]		miR-29a	0.052	[0.047, 0.194]
	miR-141	0.052	[0.026, 0.139]		miR-32	0.088	[0.032, 0.119]
	miR-145	0.072	[0.038, 0.189]		miR-16	0.059	[0.051, 0.302]
	miR-15b	0.067	[0.048, 0.230]	TWIST1	miR-195	0.173	[0.032, 0.145]
	miR-200b	0.082	[0.028, 0.134]		miR-23b	0.097	[0.044, 0.234]
	miR-200c	0.125	[0.028, 0.131]		miR-29a	0.089	[0.068, 0.355]
	miR-21	0.054	[0.047, 0.254]		miR-141	0.073	[0.023, 0.113]
	miR-25	0.078	[0.061, 0.302]		miR-145	0.104	[0.026, 0.132]
	miR-7	0.078	[0.035, 0.182]		miR-200a	0.058	[0.023, 0.118]
CDKN2A	miR-1	0.231	[0.029, 0.102]		miR-200b	0.193	[0.028, 0.111]
	miR-100	0.087	[0.041, 0.169]		miR-200c	0.390	[0.028, 0.096]

Table 5: The inferred statistical significance and active credible intervals using non-negative Bayesian LASSO for the MCC data. Only the miRNAs with significance greater than 0.05 are shown, where the bolded miRNAs are experimentally validated for repressing their targets.

mRNA	miRNA	Significance	Credible Interval	mRNA	miRNA	Significance	Credible Interval
EEF1A2	let-7f	0.631	[0.098, 0.239]	HOXD10	miR-7	0.084	[0.026, 0.098]
FSCN1	miR-133a	0.387	[0.067, 0.201]	CDKN2A	miR-1	0.501	[0.039, 0.109]
	miR-145	0.088	[0.044, 0.175]		miR-100	0.077	[0.038, 0.137]
	miR-29c	0.252	[0.095, 0.331]		miR-125b	0.131	[0.048, 0.165]
BRAF	miR-145	0.120	[0.040, 0.138]		miR-195	0.069	[0.038, 0.144]
	miR-192	0.063	[0.015, 0.058]		miR-99a	0.104	[0.041, 0.141]
	miR-221	0.072	[0.043, 0.159]	CEACAM5	miR-143	0.072	[0.062, 0.237]
	miR-223	0.337	[0.063, 0.194]		miR-145	0.073	[0.068, 0.258]
	miR-9	0.086	[0.041, 0.151]		miR-221	0.074	[0.086, 0.323]
CCND1	miR-1	0.181	[0.034, 0.117]	IGFBP6	miR-181b	0.079	[0.047, 0.176]
	miR-130a	0.100	[0.063, 0.238]	PIGR	miR-125b	0.238	[0.051, 0.170]
	miR-130b	0.126	[0.142, 0.490]	PLK1	let-7b	0.213	[0.101, 0.336]
	miR-194	0.211	[0.033, 0.109]		miR-100	0.533	[0.048, 0.132]
	miR-7	0.166	[0.054, 0.180]		miR-99a	0.161	[0.076, 0.239]
HOXD10	let-7b	0.075	[0.097, 0.362]		miR-99b	0.161	[0.076, 0.239]
	miR-103	0.080	[0.058, 0.216]	BAK1	miR-125b	0.467	[0.037, 0.106]
	let-107	0.064	[0.054, 0.205]		miR-16	0.378	[0.056, 0.168]
	miR-125b	0.110	[0.036, 0.128]		miR-25	0.087	[0.035, 0.129]
	miR-141	0.071	[0.024, 0.092]		miR-32	0.070	[0.029, 0.112]
	miR-145	0.087	[0.034, 0.122]	E2F1	miR-16	0.057	[0.042, 0.160]
	miR-15b	0.093	[0.050, 0.179]		miR-195	0.58	[0.048, 0.129]
	miR-195	0.050	[0.029, 0.118]		miR-23b	0.091	[0.039, 0.140]
	miR-200b	0.098	[0.026, 0.095]		miR-29a	0.067	[0.051, 0.190]
	miR-200c	0.219	[0.029, 0.099]	TWIST1	miR-200c	0.691	[0.032, 0.086]
	miR-21	0.057	[0.034, 0.130]		miR-145	0.104	[0.018, 0.068]
	miR-25	0.099	[0.065, 0.232]		miR-200b	0.153	[0.028, 0.083]

Table 6: The inferred statistical significance and active credible intervals using Bayesian LASSO for the MCC data. Only the miRNAs with significance greater than 0.05 are shown, where the bolded miRNAs are experimentally validated for repressing their targets.

Data	Nos of Validated Targets	LASSO	nLASSO	BLASSO	nBLASSO
MCC	23	17	13	16	20
Broad	76	34	43	47	47
Madison	99	37	52	60	84
MSKCC	106	52	52	58	59

Table 7: The numbers of experimentally validated interactions estimated by LASSO, nLASSO, BLASSO and nBLASSO.